

[Advanced Search](#) [Preferences](#) [Language Tools](#) [Search Tips](#)

web crawl dynamic search abstracts

Google Search

[Web](#) [Images](#) [Groups](#) [Directory](#)Searched the web for **web crawl dynamic search abstracts** : Results 1 - 10 of about 330. Search took 0.16 s

### FASTest Kid on the Block - Web-Ignite Corporation

... difficult to **crawl**. It allows ... results on **Web** pages, patent **abstracts**, newswires, meeting ... portals and **search** engines to ... and manage **dynamic Web** content ...

[www.web-ignite.com/editorial/indexing\\_the\\_deep\\_web.html](http://www.web-ignite.com/editorial/indexing_the_deep_web.html) - 24k - 15 Apr 2002 - [Cached](#) - [Similar pages](#)

### Spinning a Web Search : Trends in Information Retrieval

... **Web** is decentralized, **dynamic** and diverse ... let the owners **crawl** under the ... on the **Web** incorporate a ... Natural language **Search** mechanisms. ... lines of **abstracts** Keyword ...

[www.library.ucsb.edu/untangle/lager.html](http://www.library.ucsb.edu/untangle/lager.html) - 33k - [Cached](#) - [Similar pages](#)

### Scirus - About Us

... usually only **crawl** two or three ... addition to **Web** pages, Scirus ... million Beilstein **abstracts**; Over 12 ... static and **dynamic** rank values ... over other **search** engines like ...

[www.scirus.com/about/](http://www.scirus.com/about/) - 25k - 15 Apr 2002 - [Cached](#) - [Similar pages](#)

### Understanding Search Engines

... whose aim is **dynamic** updates (up to ... Who comes to **crawl** at your **web** ... pages. If a **web** crawler were to ... to your **search** query. ... index titles and **abstracts** rather than ...

[www.cudenver.edu/~mryder/help\\_data/search.html](http://www.cudenver.edu/~mryder/help_data/search.html) - 9k - [Cached](#) - [Similar pages](#)

### Search/Web

... Well indexed **Abstracts** Examples: Dialog, Lexis ... the "invisible" **web**". ... how to **crawl** frames-based ... in some **search** engines by ... Problem Two: **Dynamic** Pages. "Dynamically ...

[deskreference.org/website/searchweb1.htm](http://deskreference.org/website/searchweb1.htm) - 56k - [Cached](#) - [Similar pages](#)

### [PDF] Integrating Access to Digital Collections: MAP from Innovative ...

File Format: PDF/Adobe Acrobat - [View as HTML](#)

... **search** in Sociological **Abstracts**, finds a citation ... It does not **crawl Web** sites and ... to deliver integrated **search** results. The ... delivered in **dynamic** HTML and ...

[www.iii.com/pdf/map\\_wp.pdf](http://www.iii.com/pdf/map_wp.pdf) - [Similar pages](#)

### Search Product Ratings - Survey Results - September 1999 - ...

... indexing terms and **abstracts** which aid retrieval ... Excellent **search** in **dynamic** pages, Does not ... site" and "the **web**". ... Xavatoria + Fluid **Search**, 4, few ... fast to **crawl**. - ...

[www.searchtools.com/surveys/survey02/survey-ratings-02.html](http://www.searchtools.com/surveys/survey02/survey-ratings-02.html) - 40k - [Cached](#) - [Similar pages](#)

### Search SPIE Papers - Publications - SPIE Web

... of Contents [Hide **Abstracts**] To view ... approach of **dynamic** priority inheritance ... to current **search** engines that ... program will '**crawl**' the **Web**, finding images ...

[www.spie.org/scripts/toc.pl?volume=4312&journal=SPIE&q=spie&ab](http://www.spie.org/scripts/toc.pl?volume=4312&journal=SPIE&q=spie&ab) - 47k - [Cached](#) - [Similar pages](#)

### [PDF] Microsoft PowerPoint - 02-webcrawling

File Format: PDF/Adobe Acrobat - [View as HTML](#)

... a "broader **crawl**", because quickly ... cscf internal **search** engine User ... cf/public/**abstracts**/ # too many ... project URL, ... - **Web** site ... session ids, **dynamic** paths ...

[ciir.cs.umass.edu/cmppsci546/Slides/02-webcrawling.pdf](http://ciir.cs.umass.edu/cmppsci546/Slides/02-webcrawling.pdf) - [Similar pages](#)

## An Efficient Algorithm to Rank **Web** Resources

... in quality, (4) **dynamic**; (5) distributed and ... very helpful to **crawl** the **Web** ... of a **search** engine, it ... in ranking **Web** resources. Section 3 ... Section 4 **abstracts** a user ...  
www9.org/w9cdrom/251/251.html - 32k - [Cached](#) - [Similar pages](#)

Google

Result Page:    [1](#) [2](#) [3](#) [4](#) [5](#) [6](#) [7](#) [8](#) [9](#) [10](#)    [Next](#)

web crawl dynamic search abstracts

Google Search

[Search within results](#)

Dissatisfied with your results? [Help us improve.](#)

Try your query on: [AltaVista](#) [Excite](#) [Lycos](#) [Yahoo!](#)

[Google Home](#) - [Advertise with Us](#) - [Search Solutions](#) - [News and Resources](#) - [Language Tools](#) - [Jobs](#) - [Press](#) - [Cool Stuff...](#)

©2002 Google

Home Top: World Wide Web: Search Engines [Agents Electronic  
Commerce Metasearch Search Engines]

Change ordering: Authority Hubs (tutorials) Date Expected authority Show titles only  
**Ordered by the number of citations**

This directory is created automatically and some papers may be mislabeled. Only document within the ResearchIndex database are listed. The directory is intended to provide entry points for browsing the database and is not intended to be authoritative. Papers may not appear in all relevant categories. For example, papers in a sub-category may not appear in higher level categories.

174 Querying the World Wide Web - Mendelzon, Mihaila, Milo (1997) (Correct)

The World Wide Web is a large, heterogeneous, distributed collection of documents connected by hypertext links. The most common technology currently used for searching the Web depends on sending infor... / ... Querying the World Wide Web Alberto O. Mendelzon George A.... / ...a uniform interface to multiple search engines such as Multisurf HGN ...

120 Authoritative Sources in a Hyperlinked Environment - Kleinberg (1997) (Correct)

The network structure of a hyperlinked environment can be a rich source of information about the content of the environment, provided we have effective means for understanding it. We develop a set o... / ... variety of contexts on the World Wide Web. The central issue we address within... / ...In particular consider that current search engines typically index a sizable portion of...

119 The Anatomy of a Large-Scale Hypertextual Web Search Engine - Brin, Page (1998) (Correct)

In this paper, we present Google, a prototype of a large-scale search engine which makes heavy use of the structure present in hypertext. Google is designed to crawl and index the Web efficiently and ... / ... Anatomy of a Large-Scale Hypertextual Web Search Engine Sergey Brin and...

95 WebWatcher: A Tour Guide for the World Wide Web - Joachims, Freitag, Mitchell (1996) (Correct)

We explore the notion of a tour guide software agent for assisting users browsing the world wide web. A web tour guide agent provides assistance similar to that provided by a human tour guide in a mus... / ... WebWatcher A Tour Guide for the World...

90 Extracting Semistructured Information from the Web - Hammer, Garcia-Molina, Cho, Aranha.. (1997) (Correct)

We describe a configurable tool for extracting semistructured data from a set of HTML pages and for converting the extracted information into database objects. The input to the extractor is a declarat... / ...Semistructured Information from the Web J. Hammer H. Garcia-Molina J.... / ...a browser. Some sites do provide search engines but their query facilities are...

86 W3QS: A Query System for the World-Wide Web - David Konopnicki (1995) (Correct)

The World-Wide Web (WWW) is an ever growing, distributed, non-administered, global information resource. It resides on the worldwide computer network and allows access to heterogeneous information: te... / ...A Query System for the World-Wide Web David Konopnicki ... / ...which are built by automatic search engines called knowbots or robots . We ...

67 Database Techniques for the World-Wide Web: A Survey - Florescu, Levy, Mendelzon (1998) (Correct)

url: www.../abstr1.html] [label: Abstract, url: www.../abstr2.html] [label: Full version, url: www.../paper2.ps.Z] [label: Full version, url: www.../paper13.ps.Z] [label: Full version, url:... / ... Database Techniques for the World-Wide Web A Survey Daniela Florescu ...

61 Characterizing Browsing Strategies in the World-Wide Web - Lara Catledge (1995) (Correct)

This paper presents the results of a study conducted at Georgia Institute of Technology that captured client-side user events of NCSA's XMosaic. Actual user behavior, as determined from clientside log... / ... Browsing Strategies in the World-Wide

Web Lara D. Catledge James... / ...and there are World-Wide Web search engines available. Supporting browsing ...

**60 Searching the World Wide Web - Lawrence, Giles (1998) (Correct)**

The coverage and recency of the major World Wide Web search engines was analyzed, yielding some surprising results. The coverage of any one engine is significantly limited: No single engine indexes mo... / ... Searching the World Wide Web Steve Lawrence and C. Lee Giles ... / ...recency of the major World Wide Web search engines was analyzed yielding some...

**55 Learning to Extract Symbolic Knowledge from the World Wide Web - Craven, DiPasquo, Freitag, McCallum, .. (1998) (Correct)**

The World Wide Web is a vast source of information accessible to computers, but understandable only to humans. The goal of the research described here is to automatically create a computer understanda... / ... Symbolic Knowledge from the World Wide Web Mark Craven y Dan DiPasquo ... / ... than current keyword-based search engines. Going a step further it would...

**51 Hierarchically classifying documents using very few words - Koller, Sahami (1997) (Correct)**

The proliferation of topic hierarchies for text documents has resulted in a need for tools that automatically classify new documents within such hierarchies. Existing classification schemes which igno... / ...the contents of the World Wide Web. The bottleneck in these... / ... have been used in several internet search engines such as Yahoo Yahoo or...

**50 Building Secure and Reliable Network Applications - Birman (1996) (Correct)**

ly, the remote procedure call problem, which an RPC protocol undertakes to solve, consists of emulating LPC using message passing. LPC has a number of "properties" -- a single procedure invocation res... / ... Readings PART II THE WORLD WIDE WEB . THE WORLD WIDE WEB ...

**49 Querying Documents in Object Databases - Abiteboul, Cluet, Christophides.. (1997) (Correct)**

We consider the problem of storing and accessing documents (SGML and HTML, in particular) using database technology. To specify the database image of documents, we use structuring schemas that cons... / ...family of query languages for the Web W QL KS WebSQL MMM WebLog...

**45 Microkernels Meet Recursive Virtual Machines - Ford, Hibler, Lepreau, Tullmann.. (1996) (Correct)**

This paper describes a novel approach to providing modular and extensible operating system functionality, and encapsulated environments, based on a synthesis of microkernel and virtual machine concept... / ...virtual machines. With the advent of Web-based executable content security...

**45 Using Smart Clients to Build Scalable Services - Chad Yoshikawa (1997) (Correct)**

Individual machines are no longer sufficient to handle the offered load to many Internet sites. To use multiple machines for scalable performance, load balancing, fault transparency, and backward comp... / ... The explosive growth of the World Wide Web is straining the architecture of... / ...services such as FTP HTTP and search engines are universally accessed through...

**42 InfoSleuth: Agent-Based Semantic Integration of Information in Open.. - Bayardo, Jr., Bohrer, Brice.. (1997) (Correct)**

The goal of the InfoSleuth project at MCC is to exploit and synthesize new technologies into a unified system that retrieves and processes information in an ever-changing network of information source... / ... as internetworking and the World Wide Web have significantly expanded the... / ...Web technologies based on keyword search engines are scalable but unlike federated...

**40 Ontology-based Web Agents - Luke, Spector, Rager, Hendler (1997) (Correct)**

This paper describes SHOE, a set of Simple HTML Ontology Extensions which allow World-Wide Web authors to annotate their pages with semantic knowledge such as "I am a graduate student" or "This person... / ... Ontology-based Web Agents Sean Luke ... / ...keyword searches enabled by current search engines. We have also developed a...

**37 Queries and Computation on the Web - Serge Abiteboul (1997) (Correct)**

The paper introduces a model of the Web as an infinite, semistructured set of objects. We reconsider the classical notions of genericity and computability of queries in this new context and relate t... / ... Queries and Computation on the Web Serge Abiteboul and...

36 OBSERVER: An Approach for Query Processing in Global Information.. - Mena, Illarramendi, Kashyap, Sheth (2000) (Correct)

There has been an explosion in the types, availability and volume of data accessible in an information system, thanks to the World Wide Web (the Web) and related inter-networking technologies. In th... / ...system thanks to the World Wide Web the Web and related... / ...access as exemplified by search engines. Some also support concept based...

36 WebOS: Operating System Services for Wide Area Applications - Amin Vahdat (1997) (Correct)

In this paper, we argue for the power of providing a common set of OS services to wide area applications, including mechanisms for resource discovery, persistent storage, remote process execution, res... / ... WebOS Operating System Services for Wide ...

35 The MultiSpace: an Evolutionary Platform for Infrastructural Services - Gribble, Welsh, Brewer, Culler (1999) (Correct)

This paper presents the architecture for a Base, a clustered environment for building and executing highly available, scalable, but exible and adaptable infrastructure services. Our architecture has t... / ...collection of data repositories and web pages the Internet has become a ... / ...systems are large carrier-class web search engines portals and application-speci c web ...

33 Optimizing Regular Path Expressions Using Graph Schemas - Fernandez, Suciu (1998) (Correct)

Several languages, such as LOREL and UnQL, support querying of semi-structured data. Others, such as WebSQL and WebLog, query Web sites. All these languages model data as labeled graphs and use regula... / ... semi-structured data. Others such as WebSQL and WebLog query Web sites. All... / ...be evaluated efficiently by Web-site search engines. Introduction Query languages...

32 WebSeer: An Image Search Engine for the World Wide Web - Frankel, Swain, Athitsos (1997) (Correct)

Because of the size of the World Wide Web and its inherent lack of structure, finding what one is looking for can be a challenge. PC-Meter's March, 1996, survey found that three of the five most visit... / ... WebSeer An Image Search Engine for the ... / ... WebSeer An Image Search Engine for the World Wide Web ...

32 Web Document Clustering: A Feasibility Demonstration - Zamir, Etzioni (1998) (Correct)

Users of Web search engines are often forced to sift through the long ordered list of document "snippets" returned by the engines. The IR community has explored document clustering as an alternative ... / ... Web Document Clustering A Feasibility... / ... Abstract Users of Web search engines are often forced to sift through...

32 Semistructured and Structured Data in the Web: Going Back and Forth - Paolo Atzeni Giansalvatore (1997) (Correct)

this paper, we present the approach to the management of Web data as attacked in the Araneus unknown Semistructured and Structured Data in the Web: Going Back and Forth Paolo Atzeni, Giansalvatore M... / ...and Structured Data in the Web Going Back and Forth Paolo... / ...techniques. Since browsing and search engines present important limitations ...

31 The Michigan Internet AuctionBot: A Configurable Auction Server for.. - Wurman, Wellman, Walsh (1998) (Correct)

Market mechanisms, such as auctions, will likely represent a common interaction medium for agents on the Internet. The Michigan Internet AuctionBot is a flexible, scalable, and robust auction server t... / ...the system over the World-Wide Web as an experiment in Internet... / ...providers and specialized search engines. In addition to their use in...

31 Strong Regularities in World Wide Web Surfing - Huberman, al. (1998) (Correct)

One of the most common modes of accessing information in the World Wide Web is surfing from one document to another along hyperlinks. Several large empirical studies have revealed common patterns of s... / ... Strong Regularities in World Wide Web Surfing Bernardo A. Huberman ... / ... on the Web is through query-based search engines which enable quick access to ...

30 Searching for Images and Videos on the World-Wide Web - Smith, Chang (1996) (Correct)

We describe a prototype visual information system for searching for images and videos on the World-Wide Web. New visual information in the form of images, graphics, animations and videos is being publ... / ...Images and Videos on the World-Wide Web John R. Smith and Shih-Fu Chang ... / ...of current text-based Web search engines. The key to cataloging it is the...

**30 Applications of a Web Query Language - Gustavo Arocena University (1997) (Correct)**

In this paper we report on our experience using WebSQL, a high level declarative query language for extracting information from the Web. WebSQL takes advantage of multiple index servers without requ... / ... Applications of a Web Query Language Gustavo O. Arocena ... / ...a uniform interface to multiple search engines such as Multisurf HGN ...

**27 Web Mining: Information and Pattern Discovery on the World Wide Web - Cooley, Mobasher, Srivastava (1997) (Correct)**

Application of data mining techniques to the World Wide Web, referred to as Web mining, has been the focus of several recent research projects and papers. However, there is no established vocabulary, ... / ... Web Mining Information and Pattern...

**26 ImageRover: A Content-Based Image Browser for the World Wide Web - Sclaroff, Taycher, Cascia (1997) (Correct)**

ImageRover is a search by image content navigation tool for the world wide web. To gather images expediently, the image collection subsystem utilizes a distributed fleet of WWW robots running on diffe... / ...Image Browser for the World Wide Web Stan Sclaroff Leonid Taycher and ... / ...retrieval world wide web search engines. Introduction For a while now ...

**25 Relevance Feedback: A Power Tool for Interactive Content-Based Image.. - Yong Rui (1998) (Correct)**

Content-Based Image Retrieval (CBIR) has become one of the most active research areas in the past few years. Many visual feature representations have been explored and many systems built. While these ... / ...and the advent of the World-Wide Web there has been an explosion in the... / ... techniques in today's best text search engines such as Inquiry Alta Vista ...

**25 HPP: HTML Macro-Preprocessing to Support Dynamic Document Caching - Dougliis, Haro, Rabinovich (1997) (Correct)**

A number of techniques are available for reducing latency and bandwidth requirements for resources on the World Wide Web, including caching, compression, and delta-encoding [12]. These approaches are ... / ...for resources on the World Wide Web including caching compression and ... / ... from local CGI scripts and two popular search engines indicate that our approach promises a ...

**22 SavvySearch: A Meta-Search Engine that Learns which Search Engines to .. - Howe, Dreilinger (1997) (Correct)**

Search engines are among the most successful applications on the Web today. So many search engines have been created that it is difficult for users to know where they are, how to use them and what top... / ...most successful applications on the Web today. So many search engines have... / ... SavvySearch A Meta-Search Engine that Learns which Search Engines to...

**22 A Web-based Information System that Reasons with Structured.. - Cohen (1998) (Correct)**

The degree to which information sources are pre-processed by Web-based information systems varies greatly. In search engines like Altavista, little pre-processing is done, while in "knowledge integrat... / ... A Web-based Information System that Reasons ... / ... information systems varies greatly. In search engines like Altavista little pre-processing ...

**21 Focused crawling: a new approach to topic-specific Web resource.. - Chakrabarti, van den Berg, Dom (1999) (Correct)**

The rapid growth of the World-Wide Web poses unprecedented scaling challenges for general-purpose crawlers and search engines. In this paper we describe a new hypertext resource discovery system calle... / ...a new approach to topic-specific Web resource discovery Soumen... / ...for general-purpose crawlers and search engines. In this paper we describe a new...

**21 Inferring Web Communities from Link Topology - Gibson, Kleinberg, al. (1998) (Correct)**

The World Wide Web grows through a decentralized, almost anarchic process, and this has resulted in a large hyperlinked corpus without the kind of logical organization that can be built into more trad... / ... Inferring Web Communities from Link Topology ... / ... on the topic Harvard. Most standard search engines do not for example return...

21 Harvest: A Scalable, Customizable Discovery and Access System - Mic Bowman (1995)  
(Correct)

Rapid growth in data volume, user base, and data diversity render Internet-accessible information increasingly difficult to use effectively. In this paper we introduce Harvest, a system that provides ... / ... information on the World Wide Web gained widespread use because of ... / ... structure-preserving indexes flexible search engines and data type-specific manipulation ...

21 The PageRank Citation Ranking: Bringing Order to the Web - Page, Brin, Motwani, Winograd (1998) (Correct)

The importance of a Web page is an inherently subjective matter, which depends on the readers interests, knowledge and attitudes. But there is still much that can be said objectively about the relat... / ...Ranking Bringing Order to the Web January Abstract The...

21 ReferralWeb: Combining Social Networks and Collaborative Filtering - Kautz, Selman, Shah (1997) (Correct)

This paper appears in the Communications of the ACM, unknown ReferralWeb: Combining Social Networks and Collaborative Filtering Henry Kautz, Bart Selman, and Mehul Shah ATT Laboratories 600 Mountain ... / ...social networks on the World Wide Web. Simulation experiments we ran... / ... with the system it uses a general search engine to retrieve Web documents that...

21 CiteSeer: An Autonomous Web Agent for Automatic Retrieval and... - Bollacker, Lawrence, Giles (1998) (Correct)

Published research papers available on the World Wide Web (WWW or Web) are often poorly organized, often exist in non-text form (e.g. Postscript) documents, and increase in quantity daily. Significant... / ... CiteSeer An Autonomous Web Agent for Automatic Retrieval and... / ... a set of keywords the agent uses Web search engines and heuristics to locate and...

21 Enhanced hypertext categorization using hyperlinks - Chakrabarti, Dom, Indyk (1998) (Correct)

A major challenge in indexing unstructured hypertext databases is to automatically extract meta-data that enables structured search using topic taxonomies, circumvents keyword ambiguity, and improves ... / ... becoming increasingly important as the web is expanding search engines are... / ...important as the web is expanding search engines are proliferating and text and ...

21 Web Mining: Pattern Discovery from World Wide Web Transactions - Bamshad Mobasher (1996) (Correct)

Web-based organizations often generate and collect large volumes of data in their daily operations. Analyzing such data can help these organizations to determine the life time value of clients, design... / ... Web Mining Pattern Discovery from World...

20 Ontobroker: The Very High Idea - Fensel, Decker, Erdmann, Studer (1998) (Correct)

The World Wide Web (WWW) is currently one of the most important electronic information sources. However, its query interfaces and the provided reasoning services are rather limited. Ontobroker consist... / ... Abstract The World Wide Web WWW is currently one of the most ... / ... facilities carried out by different search engines web crawlers web indices man-made ...

20 ARACHNID: Adaptive Retrieval Agents Choosing Heuristic Neighborhoods.. - Menczer (1997) (Correct)

ARACHNID is a distributed algorithm for information discovery in large, dynamic, distributed environments such as the World Wide Web. The approach is based on a distributed, adaptive population of int... / ... environments such as the World Wide Web. The approach is based on a... / ...points provided by your favorite search engine or by browsing some digital library. ...

20 SONIA: A Service for Organizing Networked Information Autonomously - Sahami, Yusufali, Baldonado (1998) (Correct)

The recent explosion of on-line information in Digital Libraries and on the World Wide Web has given rise to a number of query-based search engines and manually constructed topical hierarchies. Howeve... / ...Libraries and on the World Wide Web has given rise to a number of... / ... given rise to a number of query-based search engines and manually constructed topical...

20 Self-Organizing Maps of Document Collections: A New Approach to.. - Lagus, al. (1996)  
(Correct)

Powerful methods for interactive exploration and search from collections of free-form textual documents are needed to manage the ever-increasing flood of digital information. In this article we presen... / ...In this article we present a method WEBSOM for automatic organization of... / ... task. Efficient search tools such as search engines have quickly emerged to aid in this...

20 Experiences with Selecting Search Engines using Meta-Search - Daniel Dreilinger (1997)  
(Correct)

Search engines are among the most useful and high profile resources on the Internet. The problem of finding information on the Internet has been replaced with the problem of knowing where search engin... / ... e-mail daniel media.mit.edu Web http www.media.mit.edu daniel ... / ... Experiences with Selecting Search Engines using Meta-Search Daniel...

19 Semistructured Data and XML - Suciu (1998) (Correct)

This paper argues that the research on semistructured data is receiving a new set of challenges with the advent of XML (Extensible Mark-up Language [Bos97, Con98]). This is a new standard approved by ... / ...and without notice e.g. data on the Web . Research on semistructured data ... / ... format. Existing Web tools browsers search engines are oriented toward document ...

19 VideoQ: An Automated Content Based Video Search System Using Visual.. - Shih-Fu Chang (1997) (Correct)

The rapidity with which digital information, particularly video, is being generated, has necessitated the development of tools for efficient search of these media. Content based visual queries have be... / ...real-time interactive system on the Web based on the visual paradigm with... / ... are needed. While there are efficient search engines for text documents today there are...

19 Analysis of a Very Large AltaVista Query Log - Craig Silverstein (1998) (Correct)

In this paper we present an analysis of a 280 GB AltaVista Search Engine query log consisting of approximately 1 billion entries for search requests over a period of six weeks. This represents approxi... / ... Our data supports the conjecture that web users differ significantly from the... / ...an analysis of a GB AltaVista Search Engine query log consisting of...

19 Merging Ranks from Heterogeneous Internet Sources - Luis Gravano (1997) (Correct)

Many sources on the Internet and elsewhere rank the objects in query results according to how well these objects match the original query. For example, a real-estate agent might rank the available hou... / ... Example Consider a World-Wide Web search engine like Excite ... / ... Example Consider a World-Wide Web search engine like Excite ...

18 Finding Related Pages in the World Wide Web - Dean, Henzinger (1999) (Correct)

When using traditional search engines, users have to formulate queries to describe their information need. This paper discusses a different approach to web searching where the input to the search proc... / ...Related Pages in the World Wide Web Jeffrey Dean Monika R.... / ... Abstract When using traditional search engines users have to formulate queries to...

18 A Machine Learning Architecture for Optimizing Web Search Engines - Boyan, Freitag, Joachims (1996) (Correct)

Indexing systems for the World Wide Web, such as Lycos and Alta Vista, play an essential role in making the Web useful and usable. These systems are based on Information Retrieval methods for indexing... / ...Learning Architecture for Optimizing Web Search Engines Justin Boyan Dayne ... / ...Architecture for Optimizing Web Search Engines Justin Boyan Dayne Freitag and...

18 Efficient Crawling Through URL Ordering - Cho, Garcia-Molina, Page (1998) (Correct)

In this paper we study in what order a crawler should visit the URLs it has seen, in order to obtain more "important" pages first. Obtaining important pages rapidly can be very useful when a crawler c... / ... when a crawler cannot visit the entire Web in a reasonable amount of time. We... / ... Web pages commonly for use by a search engine Pinkerton or a Web cache....

18 Ontobroker: Ontology based Access to Distributed and Semi-Structured.. - Decker, Erdmann, Fensel, Studer (1998) (Correct)



The World Wide Web (WWW) can be viewed as the largest multimedia database that has ever existed. However, its support for query answering and automated inference is very limited. Metadata and domain... / ... Abstract. The World Wide Web WWW can be viewed as the largest... / ... facilities carried out by different search engines web crawlers web indices man-made...

#### 18 Agents for Information Gathering - Craig Knoblock (1997) (Correct)

pear in an evolutionary fashion, driven by the market forces of applications that can benefit from using them. We believe that this bottom-up approach can lead more realistically to the development of... / ... or topic directories in the World Wide Web provide limited capabilities for... / ... scale well. Current sytems such as search engines or topic directories in the World...

#### 17 Query Decomposition and View Maintenance for Query Languages for.. - Dan Suciu (1996) (Correct)

Recently, several query languages have been proposed for querying information sources whose data is not constrained by a schema, or whose schema is unknown. Examples include: LOREL (for querying data ... / ... W QS for querying the World Wide Web and UnQL for querying... / ... by the fact that today's web search engines are restricted to content based ...

#### 17 Flexible Double Auctions for Electronic Commerce: Theory and.. - Wurman, Walsh, Wellman (1998) (Correct)

We consider a general family of auction mechanisms that admit multiple buyers and sellers, and determine market-clearing prices. We analyze the economic incentives facing participants in such auctio... / ... of online auctions on the World-Wide Web is evidence that explicit... / ... Resource finding is facilitated by search engines and shopping agents and...

#### 16 Collaborative Browsing in the World Wide Web - Gabriel Sidler (1997) (Correct)

The World Wide Web (WWW) is today the most successful service of the Internet. The richness of information available combined with easy access to this information makes it a premier information gather... / ... Browsing in the World Wide Web Gabriel Sidler... / ... people can submit keywords to search engines and hope to get a significant answer...

#### 16 Learning to Understand Information on the Internet: An Example-Based.. - Perkowitz, Doorenbos, Etzioni, Weld (Correct)

The explosive growth of the Web has made intelligent software assistants increasingly necessary for ordinary computer users. Both traditional approaches --- search engines, hierarchical indices --- ... / ... Abstract. The explosive growth of the Web has made intelligent software... / ... users. Both traditional approaches search engines hierarchical indices and...

#### 16 GLOSS: Text-Source Discovery over the Internet - Luis Gravano Columbia (1999) (Correct)

74477422 487247 Thesaurus 11382655 3695 Conference 7246145 11934 Organization 9374199 62051 Class 4211136 2962 Numbers (ISBN, ...) 2445828 12637 Report Numbers 7833 7508 Totals 130,340,123 1,08... / ... all documents. Some systems e.g. Web search engines discard the documents ... / ... categories single versus distributed search engines. A single search engine builds a...

#### 16 Newsgroup Exploration with WEBSOM Method and Browsing Interface - Timo Honkela (1996) (Correct)

The current availability of large collections of full-text documents in electronic form emphasizes the need for intelligent information retrieval techniques. Especially in the rapidly growing World Wi... / ... Newsgroup Exploration with WEBSOM Method and Browsing Interface ... / ... become necessary. Efficient search engines have been developed to aid in the...

#### 15 Revisitation Patterns in World Wide Web Navigation - Tauscher, Greenberg (1997) (Correct)

We report on users' revisitation patterns to World Wide Web pages, and use these to lay an empirical foundation for the design of history mechanisms in web browsers. Through history, a user can return... / ... Revisitation Patterns in World Wide Web Navigation Linda Tauscher and Saul... / ... decrease resource use by supplanting search engines for finding old pages and by...

#### 15 A Network Architecture for Heterogeneous Mobile Computing - Brewer, Amir, Balakrishnan.. (1998) (Correct)

This paper summarizes the results of the BARWAN project, which focused on enabling truly useful mobile networking across an extremely wide variety of real-world networks and mobile devices. We present... / ... services such as search engines

and web access. This is an alternative to the ... / ... directions to global services such as search engines and web access. This is an...

#### 14 WebMate : A Personal Agent for Browsing and Searching - Liren Chen (1998) (Correct)

The World-Wide Web is developing very fast. Currently, finding useful information on the Web is a time consuming process. In this paper, we present WebMate, an agent that helps users to effectively br... / ... WebMate A Personal Agent for Browsing... / ... to them by sending a query to a search engine such as Altavista by following ...

#### 14 Finding Salient Features for Personal Web Page Categories - Wulfekuhler, Punch (1997) (Correct)

We examine techniques that "discover" features in sets of pre--categorized documents, such that similar documents can be found on the World Wide Web. First, we examine techniques which will classify t... / ... Finding Salient Features for Personal Web Page Categories Marilyn R.... / ...is why identical queries to different search engines produce different results. Query...

#### 14 Visual Information Retrieval from Large Distributed On-line.. - Chang (1997) (Correct)

ion --- Images may be indexed at various levels, including feature (e.g., color, texture, and shape), object (e.g., moving foreground object), syntax (e.g., video shot), and semantics (e.g., image sub... / ...retrieval system VIRS in the Webbased environment we present our... / ...photo stocks and WorldWide Web search engines. A high-level taxonomy will be...

#### 14 Image Digestion and Relevance Feedback in the ImageRover WWW Search.. - Taycher, Cascia, Sclaroff (1997) (Correct)

ImageRover is a search by image content navigation tool for the world wide web. The staggering size of the WWW dictates certain strategies and algorithms for image collection, digestion, indexing, and... / ...navigation tool for the world wide web. The staggering size of the WWW... / ...Feedback in the ImageRover WWW Search Engine Leonid Taycher Marco La Cascia ...

#### 13 A fully automated content based video search engine supporting.. - Chang, Chen, Meng, Sundaram, Di Zhong (1997) (Correct)

The rapidity with which digital information, particularly video, is being generated, has necessitated the development of tools for efficient search of these media. Content based visual queries have be... / ...a novel interactive system on the Web based on the visual paradigm with... / ... A Fully Automated Content Based Video Search Engine Supporting Spatio-Temporal Queries ...

#### 13 Context and Page Analysis for Improved Web Search - Lawrence, Giles (1998) (Correct)

NEC Research Institute has developed a metasearch engine that improves the efficiency of Web searches by downloading and analyzing each document and then displaying results that show the query terms i... / ...that improves the efficiency of Web searches by downloading and... / ... Several popular and useful search engines-such as AltaVista Excite HotBot ...

#### 13 A Fully Automated Content-Based Video Search Engine Supporting.. - Shih-Fu Chang Member (1998) (Correct)

The rapidity with which digital information, particularly video, is being generated has necessitated the development of tools for efficient search of these media. Content-based visual queries have bee... / ...a novel interactive system on the Web based on the visual paradigm with... / ... A Fully Automated Content-Based Video Search Engine Supporting Spatiotemporal Queries ...

#### 13 Cluster-based Language Models For Distributed Retrieval - Jinxi Xu And (1999) (Correct)

Effective retrieval in a distributed environment is an important but difficult problem. Lack of effectiveness appears to have three causes. First, collection selection based on word histograms is not ... / ... more so in the future. The World Wide Web for example already consists of... / ...as meta searching but current meta search engines such as MetaCrawler typically...

#### 12 Trawling the web for emerging cyber-communities - Ravi Kumar Prabhakar (1999) (Correct)

The web harbors a large number of communities - groups of content-creators sharing a common interest which manifests itself as a set of web pages. Whereas newgroups and commercial web directories to... / ... Trawling the web for emerging cyber-communities ...

#### 12 Relevance Feedback Techniques in Interactive Content-Based Image.. - Rui, Huang, Mehrotra

**(Correct)**

Content-Based Image Retrieval (CBIR) has become one of the most active research areas in the past few years. Many visual feature representations have been explored and many systems built. While these ... / ...and the advent of the World-Wide Web there has been an explosion in the... / ...techniques in today's best text search engines such as Yahoo Alta Vista Lycos ...

**12 Processing Queries for First-Few Answers - Bayardo, Jr. (1996) (Correct)**

Special support for quickly finding the first-few answers of a query is already appearing in commercial database systems. This support is useful in active databases, when dealing with potentially unma... / ... including those for the world wide web provide functionality for lazily... / ...result is simply too costly. Various search engines including those for the world wide ...

**12 Fast and Intuitive Clustering of Web Documents - Oren Zamir (1997) (Correct)**

Conventional document retrieval systems (e.g., Alta Vista) return long lists of ranked documents in response to user queries. Recently, document clustering has been put forth as an alternative method ... / ... Fast and Intuitive Clustering of Web Documents Oren Zamir and Oren ... / ...of snippets returned from Web search engines. First we show that ...

**12 Improving Automatic Query Expansion - Mandar Mitra (Correct)**

Most casual users of IR systems type short queries. Recent research has shown that adding new words to these queries via blind feedback, without any input from the user, improves the performance of su... / ...the proliferation of the World Wide Web and the widespread use of Web search... / ... Wide Web and the widespread use of Web search engines the number of casual users of IR...

**12 WebSuite---A Tool Suite For Harnessing Web Data - Catriel Beeri (1998) (Correct)**

We present a system for searching, collecting, and integrating Web-resident data. The system consists of five tools, where each tool provides a specific functionality aimed at solving one aspect of th... / ... email oshmu CS.Technion.AC.IL WebSuite-A Tool Suite For Harnessing Web ... / ... filtering and selection than most Web search engines e.g. . This powerful...

**12 Techniques for Developing and Measuring High Performance Web Servers.. - James Hu (Correct)**

High-performance Web servers are essential to meet the growing demands of the Internet and large-scale intranets. Satisfying these demands requires a thorough understanding of key factors affecting We... / ...and Measuring High Performance Web Servers over High Speed Networks ...

**12 Using Reinforcement Learning to Spider the Web Efficiently - Rennie, McCallum (1999) (Correct)**

Consider the task of exploring the Web in order to find pages of a particular kind or on a particular topic. This task arises in the construction of search engines and Web knowledge bases. This paper ... / ...Reinforcement Learning to Spider the Web Efficiently Jason Rennie ... / ...task arises in the construction of search engines and Web knowledge bases. This paper...

**12 Techniques for Developing and Measuring High-Performance Web Servers.. - James Hu (Correct)**

High-performance Web servers are essential to meet the growing demands of the Internet and large-scale intranets. Satisfying these demands requires a thorough understanding of key factors affecting We... / ...and Measuring High-Performance Web Servers over ATM Networks James C. ...

**11 Document Categorization and Query Generation on the World Wide Web.. - Boley (1999) (Correct)**

We present WebACE, an agent for exploring and categorizing documents on the World Wide Web based on a user profile. The heart of the agent is an unsupervised categorization of a set of documents, comb... / ...Query Generation on the World Wide Web Using WebACE Daniel Boley Maria... / ...continues to grow rapidly. Powerful search engines have been developed to aid in...

**11 Combining Textual and Visual Cues for Content-based Image Retrieval.. - Cascia, Sethi, Sclaroff (1998) (Correct)**

A system is proposed that combines textual and visual statistics in a single index vector for content-based search of a WWW

image database. Textual statistics are captured in vector form using latent ... / ...Image Retrieval on the World Wide Web  
Marco La Cascia Saratendu Sethi ... / ... led to the birth of a number of image search engines . The web's staggering ...

# 11 Synchronizing a database to Improve Freshness - Cho, Garcia-Molina (2000) (Correct)

In this paper we study how to refresh a local copy of an autonomous data source to maintain the copy up-to-date. As the size of the data grows, it becomes more difficult to maintain the copy "fresh," mainly based on data collected from web sites for more than months and we... / ...for local analysis. Similarly a web search engine copies portions of the web and then...

# 11 Interactive Query and Search in Semistructured Databases - Roy Goldman (1998) (Correct)

Semistructured graph-based databases have been proposed as well-suited stores for World-Wide Web data. Yet so far, languages for querying such data are too complex for casual Web users. Further, proposed... / ...as well-suited stores for World-Wide Web data. Yet so far languages for... / ...For searching the entire Web search engines are a well-proven successful...

# 11 Data Management for XML: Research Directions - Widom (1999) (Correct)

This paper is a July 1999 snapshot of a "whitepaper" that I've been working on. The purpose of the whitepaper, which I initially drafted in April 1999, was to formulate and put into prose my thoughts ... / ...the potential impact is significant Web servers and applications encoding... / ... keyword-based searches as provided by search engines for example that understand ...

# 11 Information Gathering in the World-Wide Web: The W3QL Query Language.. - Konopnicki, Shmueli (1998) (Correct)

ing with credit is permitted. To copy otherwise, to republish, to post on servers, to redistribute to lists, or to use any component of this work in other works, requires prior specific permission and... / ...Gathering in the World-Wide Web The W QL Query Language and the... / ... Yahoo and others. These sites employ search engines known as robots or knowbots ...

# 11 Subquadratic Approximation Algorithms For Clustering Problems in High .. - Allan Borodin (Correct)

One of the central problems in information retrieval, data mining, computational biology, statistical analysis, computer vision, geographic analysis, pattern recognition, distributed protocols is the ... / ... pattern recognition biology web search engines distributed... / ...example in the Alta Vista tm search engine for the clustering defined by ...

# 10 Dienst: Building a Production Technical Report Server - James Davis (1995) (Correct)

Dienst is a protocol and implementation that provides Internet access to a distributed, decentralized multi-format document collection. It supports full text and boolean searches, thumbnail visual bro... / ...Using publicly available World Wide Web clients users may search the... / ...interoperability among heterogeneous search engines reliability logical document...

# 10 Digital Libraries and Autonomous Citation Indexing - Lawrence, Giles, Bollacker (1999) (Correct)

ificant improvements or criticisms of earlier work, and . helping limit the wasteful duplication of prior research. Citation indices can also be used to analyze research trends, identify emerging ... / ...libraries has passed. The Web promises to make more scientific... / ...quickly can be difficult because Web search engines have difficulty keeping upto date ...

# 10 Combining Collaborative Filtering with Personal Agents for Better.. - Nathaniel Good Ben (1999) (Correct)

Information filtering agents and collaborative filtering both attempt to alleviate information overload by identifying which items a user will find worthwhile. Information filtering (IF) focuses on... / ... more and more books journal articles web pages and movies are created. As... / ... Belkin and Croft Internet search engines are popular IR systems and the...

# 10 Grouper: A Dynamic Clustering Interface to Web Search Results - Zamir, Etzioni (1999) (Correct)

Users of Web search engines are often forced to sift through the long ordered list of document "snippets" returned by the engines. The IR community has explored document clustering as an alternative m... / ...A Dynamic Clustering Interface to Web Search Results Oren Zamir and Oren ... / ... Abstract Users of Web search engines are often forced to sift through the...

# 10 An Image and Video Search Engine for the World-Wide Web - Smith, Chang (1997) (Correct)

We describe a visual information system prototype for searching for images and videos on the WorldWide Web. New visual information in the form of images, graphics, animations and videos is being publi... / ... Video Search Engine for the World-Wide Web John R. Smith and Shih-Fu Chang ... / ... An Image and Video Search Engine for the World-Wide Web John R....

# 10 Using Agents to Improve the Usability and Usefulness of the.. - Thomas, Fischer (1996) (Correct)

The World-Wide Web (WWW) has emerged as a new type of information space. Its lack of central control mechanisms leads to many interesting new features but at the same time has the potential danger tha... / ...and Usefulness of the World-Wide Web Christoph G. Thomas ... / ...skill for example the use of search engines in the WWW is not standardized....

# 10 Machine Learning for Adaptive User Interfaces - Pat Langley (1997) (Correct)

In this paper we examine the growing interest in personalized user interfaces and explore the potential of machine learning in meeting that need. We briefly review progress in developing fielded app... / ... and search engines for the World Wide Web. Moreover there is every indication... / ...and most recently with browsers and search engines for the World Wide Web. Moreover ...

# 10 A Multi-Similarity Algebra - Adali, Bonatti, Sapino, Subrahmanian (1998) (Correct)

The need to automatically extract and classify the contents of multimedia data archives such as images, video, and text documents has led to significant work on similarity based retrieval of data. To ... / ... Example Integrating multiple web search engines Internet search... / ...methods using the Integrated Search Engine I.SEE as the testbed. ...

# 9 Storage and Retrieval of Feature Data for a Very Large Online Image.. - Chad Carson (1996) (Correct)

As network connectivity has continued its explosive growth and as storage devices have become smaller, faster, and less expensive, the number of online digitized images has increased rapidly. Successf... / ... A recent search of the World Wide Web found million pages containing the ... / ... via forms sorted lists and search engines. Image queries can rely on textual...

# 9 Patterns of Search: Analyzing and Modeling Web Query Refinement - Lau, Horvitz (1999) (Correct)

We discuss the construction of probabilistic models centering on temporal patterns of query refinement. Our analyses are derived from a large corpus of Web search queries extracted from server log... / ...of Search Analyzing and Modeling Web Query Refinement Tessa Lau ... / ...network-based services. Web-based search engines such as Excite AltaVista and Lycos...

# 9 Text / Relational Database Management Systems: Harmonizing SQL and.. - Blake Consens (1994) (Correct)

Combined text and relational database support is increasingly recognized as an emerging need of industry, spanning applications requiring text fields as parts of their data (e.g., for customer suppo... / ... process text sub-queries on full-text search engines. Introduction The application...

# 9 Embedding Knowledge in Web Documents - Martin, Eklund (1999) (Correct)

The paper argues for the use of general and intuitive knowledge representation languages (and simpler notational variants, e.g. subsets of natural languages) for indexing the content of Web documents ... / ... Embedding Knowledge in Web Documents Philippe Martin and...

# 9 Continual Queries for Internet Scale Event-Driven Information Delivery - Ling Liu (1999) (Correct)

In this paper we introduce the concept of continual queries, describe the design of a distributed event-driven continual query system \Gamma OpenCQ, and outline the initial implementation of OpenCQ on... / ... management systems such as DBMSs and Web search engines OpenCQ exhibits two... / ...systems such as DBMSs and Web search engines OpenCQ exhibits two important...

# 8 Mining the Link Structure of the World Wide Web - Chakrabarti, Dom, Gibson, Kleinberg, ... (1999) (Correct)

The World Wide Web contains an enormous amount of information, but it can be exceedingly difficult for users to locate resources that are both high in quality and relevant to their information needs. ... / ...the Link Structure of the World Wide Web Soumen Chakrabarti Byron E. ... / ... off-the-shelf fashion. Index-based search engines for the WWW have been one of the...

#### 8 MetaSEEk: A Content-Based Meta-Search Engine for Images - Mandis Beigi, Ana B. Benitez, and.. (1997) (Correct)

Search engines are the most powerful resources for finding information on the rapidly expanding World Wide Web (WWW). Finding the desired search engines and learning how to use them, however, can be v... / ...on the rapidly expanding World Wide Web WWW. Finding the desired search... / ... MetaSEEk A Content-Based Meta-Search Engine for Images Mandis Beigi Ana B....

#### 8 Overview of TREC-7 Very Large Collection Track - David Hawking (Correct)

In line with the wishes of last year's participants, this year's VLC track was essentially a re-run of last year's with a five-fold increase in data size. The data used was a completely new 100-gigaby... / ...new -gigabyte collection of Web documents the VLC whose... / ... better than several well-known search engines can be produced from queries...

#### 8 Building Domain-Specific Search Engines with Machine Learning.. - McCallum, Nigam, Rennie, Seymore (1999) (Correct)

Domain-specific search engines are becoming increasingly popular because they offer increased accuracy and extra features not possible with the general, Web-wide search engines. For example, www.camps... / ...not possible with the general Web-wide search engines. For example ... / ... Building Domain-Specific Search Engines with Machine Learning Techniques ...

#### 8 Metadata for Digital Libraries: Architecture and Design Rationale - Baldonado, Chang, Gravano, Paepcke (1997) (Correct)

In a distributed, heterogeneous, proxy-based digital library, autonomous services and collections are accessed indirectly via proxies. To facilitate metadata compatibility and interoperability in such... / ... Dialog Information Service World-Wide Web search engines automatic document... / ...Information Service World-Wide Web search engines automatic document summarizers ...

#### 8 The Internet2 Distributed Storage Infrastructure Project: An.. - Micah Beck And (1998) (Correct)

ess to certain services by applying resources to those services alone. This structure inhibits the development of a viable economic model for differential investment in new infrastructure. But the nee... / ...Because the performance of the Web is dependent on the real-time... / ...backend database systems and remote search engines. But an aggregation of such content...

#### 8 Information Forage through Adaptive Visualization - Dmitri Roussinov (1998) (Correct)

Automatically created maps of concepts improve navigation in a collection of text documents. We report our research on leveraging navigation by providing interactively the ability to modify the maps t... / ...real-time a map of concepts found in Web documents returned by a commercial... / ...documents returned by a commercial search engine. KEYWORDS Intelligent searching ...

#### 8 Overview of TREC-6 Very Large Collection Track - David Hawking (Correct)

The emergence of real world applications for text collections orders of magnitude larger than the TREC collection has motivated the introduction of a Very Large Collection track within the TREC framew... / ...scale in the commercial world and Web search engines such as HotBot claim... / ... scale in the commercial world and Web search engines such as HotBot claim to index in...

#### 7 Narcissus: Visualising Information - Hendley Drew (1995) (Correct)

It is becoming increasingly important that support is provided for users who are dealing with complex information spaces. The need is driven by the growing number of domains where there is a requireme... / ...example is the World Wide Web but other domains include software ... / ... and although there are tools such as search engines which can help users to find ...

#### 7 Mining Longest Repeating Subsequences To Predict World Wide Web.. - Pitkow, Pirolli (1999)

(Correct)

Modeling and predicting user surfing paths involves tradeoffs between model complexity and predictive accuracy. In this paper we explore predictive modeling techniques that attempt to reduce model com... / ... SUBSEQUENCES TO PREDICT WORLD WIDE WEB SURFING James Pitkow and Peter... / ... interaction. For instance the Google search engine assumes that a model of surfing...

7 Results and Challenges in Web Search Evaluation - Hawking, Craswell, Thistlewaite (1999)(Correct)

A frozen 18.5 million page snapshot of part of the Web has been created to enable and encourage meaningful and reproducible evaluation of Web search systems and techniques. This collection is being us... / ... Results and Challenges in Web Search Evaluation David Hawking ... / ...rankings produced by public Web search engines is by no means state-of-the-art. ...

7 Ontobroker: Or How to Enable Intelligent Access to the WWW - Fensel, Decker, Erdmann, Studer (1998) (Correct)

The World Wide Web (WWW) is currently one of the most important electronic information sources. However, its query interfaces and the provided reasoning services are rather limited. Ontobroker con... / ... Abstract. The World Wide Web WWW is currently one of the most... / ...facilities carried out by different search engines web crawlers web indices man-made ...

7 A Network Measurement Architecture for Adaptive Applications - Stemm, Katz, Seshan (2000)(Correct)

The quality of network connectivity between a pair of Internet hosts can vary greatly. Some hosts may communicate over high bandwidth, low latency, uncongested paths, while others communicate overmuch... / ...allows clients to download mirrored web objects within of the fastest... / ...expected transfer time. Clients using search engines could post-process the query results ...

7 Measuring Index Quality using Random Walks on the Web - Henzinger, Heydon, Mitzenmacher.. (1999) (Correct)

Recent research has studied how to measure the size of a search engine, in terms of the number of pages indexed. In this paper, we consider a different measure for search engines, namely the quality ... / ...Quality using Random Walks on the Web Monika R. Henzinger Allan Heydon... / ... studied how to measure the size of a search engine in terms of the number of pages...

7 LogicWeb: Enhancing the Web with Logic Programming - Andrew Davison (1996) (Correct)

LogicWeb is a client-side logic programming tool for the World Wide Web, which allows the Web to be viewed in a more abstract way: Web pages can be rephrased as logic programming modules, and hypertext... / ... LogicWeb Enhancing the Web with Logic Programming Andrew...

7 Adaptive Information Agents in Distributed Textual Environments - Menczer, Belew (1998)(Correct)

Hypertext environments such as the Web are rich with both word and link cues that can be exploited by autonomous agents performing distributed tasks on behalf of the user. This paper characterizes suc... / ... Hypertext environments such as the Web are rich with both word and link... / ...to your favorite digital library or search engine on the Web and received a long list ...

7 Automatic Resource list Compilation by Analyzing Hyperlink Structure.. - Chakrabarti, Dom, Gibson, Keinberg.. (1998) (Correct)

We describe the design, prototyping and evaluation of ARC, a system for automatically compiling a list of authoritative web resources on any (sufficiently broad) topic. The goal of ARC is to compile r... / ... compiling a list of authoritative web resources on any sufficiently broad ... / ...be used to re-order the output of a search engine. For a more detailed review of ...

6 Large-Scale Information Retrieval with Latent Semantic Indexing - Letsche, BERRY (1997)(Correct)

As the amount of electronic information increases, traditional lexical (or Boolean) information retrieval techniques will

become less useful. Large, heterogeneous collections will be difficult to se... / ...executing. In addition a World-Wide Web interface was created to allow... / ...to build both serial and distributed search engines. In addition Section presents a...

#### 6 A Layered Approach To Nlp-Based Information Retrieval - Flank (1998) (Correct)

A layered approach to information retrieval permits the inclusion of multiple search engines as well as multiple databases, with a natural language layer to convert English queries for use by the vari... / ...PNI now operates on the World Wide Web www.publishersdepot.com ... / ... permits the inclusion of multiple search engines as well as multiple databases with...

#### 6 Supporting Social Navigation on the World Wide Web - Dieberger (1997) (Correct)

This paper discusses a navigation behavior on Internet information services, in particular the World Wide Web, which is characterized by pointing out of information using various communication tools. ... / ...Social Navigation on the World Wide Web Andreas Dieberger Georgia...

#### 6 Wrapper Generation for Web Accessible Data Sources - Jean-Robert Gruser (1998) (Correct)

There is an increase in the number of data sources that can be queried across the WWW. Such sources typically support HTML forms-based interfaces and search engines query collections of suitably index... / ... Wrapper Generation for Web Accessible Data Sources ... / ...HTML forms-based interfaces and search engines query collections of suitably...

#### 6 An Overview of Audio Information Retrieval - Jonathan Foote (1998) (Correct)

The problem of audio information retrieval is familiar to anyone who has returned from vacation to find an answering machine full of messages. While there is not yet an "AltaVista" for the audio data ... / ...familiar to many through the popular web search engines such as Lycos or... / ...to many through the popular web search engines such as Lycos or AltaVista. The...

#### 6 Cluster Reserves: A Mechanism for Resource Management in... - Aron, Druschel, Zwaenepoel (2000) (Correct)

In network (e.g., Web) servers, it is often desirable to isolate the performance of different classes of requests from each other. That is, one seeks to achieve that a certain minimal proportion of ser... / ... ABSTRACT In network e.g. Web servers it is often desirable to... / ... retrieval and electronic commerce and search engines. It is often desirable that...

#### 6 Structuring and Visualising the WWW by Generalised Similarity Analysis - Chen (1997) (Correct)

This paper describes a generic approach to structuring and visualising a hypertext-based information space on the WWW. This approach, called Generalised Similarity Analysis (GSA), provides a unifi... / ...information spaces on the World-Wide Web WWW raises a number of practical... / ...such as search results returned by search engines most hypertext documents on the WWW ...

#### 6 The Order of Things: Activity-Centred Information Access - Matthew Chalmers (1998) (Correct)

This paper focuses on the representation and access of Web-based information, and how to make such a representation adapt to the activities or interests of individuals within a community of users. The... / ...on the representation and access of Web-based information and how to make... / ...techniques and so limits the power of search engines. In contrast to traditional methods ...

#### 6 User Interactions with Everyday Applications as Context for... - Budzik, Hammond (2000) (Correct)

Our central claim is that user interactions with everyday productivity applications (e.g., word processors, Web browsers, etc.) provide rich contextual information that can be leveraged to support jus... / ...applications e.g. word processors Web browsers etc. provide rich... / ...context. Moreover a recent study of search engine queries showed that on average ...

#### 6 Automatic Web Page Categorization by Link and Context Analysis - Attardi, Gulli, Sebastiani (1999) (Correct)

Assistance in retrieving documents on the World Wide Web is provided either by search engines, through keyword-based queries, or by catalogues, which organize documents into hierarchical collections... / ... Automatic Web Page Categorization by Link and... / ...World Wide Web is provided either by search engines through keyword-based queries or...

#### 6 Learning to Query the Web - Cohen, Singer (1996) (Correct)



The World Wide Web (WWW) is filled with "resource directories"---i.e., documents that collect together links to all known documents on a specific topic. Keeping resource directories up-to-date is dif... / ... Learning to Query the Web William W. Cohen and Yoram Singer ... / ... or series of queries for a WWW search engine. This query can be used at a later...

#### 6 Modeling and Querying Semi-Structured Data - Cluet (1997) (Correct)

This paper does not develop all of them but concentrates on the problems of modeling and querying (efficiently) semi-structured data. It is by no means exhaustive. It compares various works with some ... / ... Since its creation in the Web has known a tremendous success due to ... / ...of these research projects is that search engines as found on the Web e.g. Altavista ...

#### 6 Learning to Extract Keyphrases from Text - Turney (1999) (Correct)

Many academic journals ask their authors to provide a list of about five to fifteen key words, to appear on the first page of each article. Since these key words are often phrases of two or more words... / ...be used to analyze usage patterns in web server logs. These examples motivate... / ... has a specific need. When a search engine form has a field labelled key...

#### 5 Towards a Framework for Developing Mobile Agents for Managing.. - Jonathan Dale And (1997) (Correct)

In this paper, we present a layered organisation for our agent-based applications. This organisation is derived from analysing the commonalities that are shared by our separate applications. Within this ... / ...If we consider the World Wide Web as a case study the prominent tool... / ...are text-based in nature as with search engines for the World Wide Web for example...

#### 5 Using Fagin's Algorithm for Merging Ranked Results in Multimedia.. - Edward Wimmers Laura (1998) (Correct)

A distributed multimedia information system allows applications to access a variety of data, of different modalities, stored in data sources with their own specialized search capabilities. In such a ... / ...as scientific applications on the Web as well as on the intranet. With the ... / ...These systems as well as many text search engines and other more specialized...

#### 5 Musag: an agent that learns what you mean - Claudia Goldman And (1996) (Correct)

This paper presents a system that carries out highly effective searches over collections of textual information, such as those found on the Internet. The system is comprised of two major parts. The ... / ...continuously e.g. the World Wide Web . We are investigating two... / ...two main problems with current text search engines which are largely based on...

#### 5 Matchmaking among Heterogeneous Agents on the Internet - Katia Sycara (1999) (Correct)

The Internet is not only providing data for users to browse, but also databases to query, and software agents to run. Due to the exponential increase of deployed agents on the Internet, automating the... / ... fl Application in the Web. One of the main application domains... / ...proposed a method to look for better search engines that may provide more relevant data...

#### 5 Object-based Navigation: An Intuitive Navigation Style for.. - Kyoji Hirata Sougata (1997) (Correct)

In this paper, we present the idea of object-based navigation. Object-based navigation is a navigation style based upon the characteristics at the object level, that is contents of the objects and ... / ...service systems for the World-Wide Web and have evaluated the navigational ... / ... extensibility of our tools. Multimedia search engines including COIR extract the...

#### 5 SPIRIT: Sequential Pattern Mining with Regular Expression Constraints - Minos Garofalakis Bell (1999) (Correct)

Discovering sequential patterns is an important problem in data mining with a host of application domains including medicine, telecommunications, and the World Wide Web. Conventional mining systems pr... / ... telecommunications and the World Wide Web. Conventional mining systems provide ... / ...present in today's keyword-based WWW search engines. The idea is to allow users to...

#### 5 Object-oriented and Database Concepts for the Design of Networked.. - Norbert Fuhr (1996) (Correct)

By using data abstraction concepts from database and objectoriented systems and combining them with uncertain inference, we develop a new approach for the design of information retrieval (IR) systems,... / ... are available on the network. Popular Web search engines aim at indexing almost ... / ... available on the network. Popular Web search engines aim at indexing

almost any document ...

#### 5 Using Machine Learning To Improve Information Access - Sahami (1999) (Correct)

The explosion of on-line information has given rise to many query-based search engines (such as Alta Vista) and manually constructed topic hierarchies (such as Yahoo! ). But with the current growth ra... / ...the Planet Saturn node. . A Web page dynamically generated by SONIA ... / ...has given rise to many query-based search engines such as Alta Vista and manually...

#### 5 Indexing and Retrieval of Scientific Literature - Lawrence, Bollacker, Giles (1999) (Correct)

The web has greatly improved access to scientific literature. However, scientific articles on the web are largely disorganized, with research articles being spread across archive sites, institution si... / ... Abstract The web has greatly improved access to... / ...literature and the major web search engines typically do not index the content ...

#### 5 A Mobile Agent Architecture for Distributed Information Management - Dale, DeRoure (1997) (Correct)

this paper, we present an architecture that supports mobile agents which carry out distributed information management tasks to help the user discover, create, maintain and integrate information which ... / ...If we consider the World Wide Web WWW as a case study the prominent... / ...are text-based in nature as with search engines for the WWW. These can deal with text ...

#### 5 A Layered Architecture for Querying Dynamic Web Content - Davulcu, Freire, Kifer, Ramakrishnan (1999) (Correct)

The design of webbases, database systems for supporting Webbased applications, is currently an active area of research. In this paper, we propose a 3-layer architecture for designing and implementing ... / ...Architecture for Querying Dynamic Web Content Hasan Davulcu ...

#### 5 On Caching Search Engine Results - Evangelos Markatos (1999) (Correct)

In this paper we explore the problem of Caching of Search Engine Query Results in order to reduce the computing and I/O requirements needed to support the functionality of a search engine of the world... / ...of a search engine of the world-wide web. Based on traces from search engines ... / ... On Caching Search Engine Results Evangelos P. Markatos ...

#### 5 Keeping Up With The Changing Web - Brewington, Cybenko (2000) (Correct)

Our access to information today is unprecedented in history. However, information depreciates in value as it gets older, and the problem of updating information to keep it current presents new design... / ... Keeping Up With The Changing Web Brian E. Brewington George... / ... Web. We quantify what it means to for search engines to be up-to-date and estimate how...

#### 5 A Conceptual Graph Model for W3C Resource Description Framework - Corby, Dieng, Hébert (2000) (Correct)

With the aim of building a "Semantic Web", the content of the documents must be explicitly represented through metadata in order to enable contents-guided search. Our approach is to exploit a stan... / ... With the aim of building a Semantic Web the content of the documents must... / ... world. But the existing keyword-based search engines do not take into account the...

#### 4 A non-invasive learning approach to building web user profiles - Chan (1999) (Correct)

Introduction Recently researchers have started to make web browsers more adaptive and personalized. A personalized web browser caters to the user's interests and an adaptive one learns from the users... / ... KDD- Workshop on Web Usage Analysis and User Profiling to ... / ...personaized on-line search our search engine consults multiple existing search...

#### 4 Conceptual Views over the Web - Tiziana Catarci (1997) (Correct)

The Internet has made available an enormous quantity of information to a disparate variety of people. The amount of information, the typical access modality (i.e. browsing), and the open growth of the... / ... Conceptual Views over the Web Tiziana Catarci Luca Iocchi ... / ...mechanism. Popular keyword-based search engines can be regarded as first generation...

#### 4 Web Page Categorization and Feature Selection Using Association Rule.. - Jerome Moore Eui-Hong (1997) (Correct)

Clustering techniques have been used by many intelligent software agents in order to retrieve, #lter, and categorize documents available on the World Wide Web. Clustering is also useful in extracting... / ... Web Page Categorization and Feature... / ...continues to grow rapidly. Powerful search engines have been developed to aid in...

#### 4 An IR Approach for Translating New Words from Nonparallel, Comparable ... - Pascale Fung (Correct)

this paper, we describe a new method which combines IR and NLP techniques to extract new word translation from automatically downloaded English-Chinese nonparallel newspaper texts. unknown An IR Appro... / ...repository known as the World Wide Web. Various traditional information... / ...enable efficient access of the WWW-search engines indexing relevance feedback query ...

#### 4 A Semidiscrete Matrix Decomposition for Latent Semantic Indexing in... - Tamara Kolda Oak (1997) (Correct)

this article we propose replacing the SVD with the semidiscrete decomposition (SDD). We will describe the SDD approximation, show how to compute it, and compare the SDD-based LSI method to the SVD-bas... / ... and search engines for the World Wide Web such as Alta Vista. Oftentimes ... / ... string-matching tool in Unix and search engines for the World Wide Web such as Alta...

#### 4 Materializing the Web - De Rosa, Catarci, Iocchi, Nardi.. (1998) (Correct)

In this paper we present a novel approach to accessing the Web, that enables for automatically acquiring data from Web sites and making them accessible to the user through a database query paradigm. T... / ... Materializing the Web Mattia De Rosa Tiziana Catarci ... / ... of interest. Popular keyword-based search engines can be considered as GIMSS where...

#### 4 Searching the Web with SHOE - Heflin, Hendler (2000) (Correct)

Although search engine technology has improved in recent years, there are still many types of searches that return unsatisfactory results. This situation can be greatly improved if web pages use a ... / ... Searching the Web with SHOE Jeff Heflin and James... / ... Abstract Although search engine technology has improved in recent ...

#### 4 The Stanford InfoBus and Its Service Layers - Augmenting the Internet .. - Röscheisen, Baldonado, Chang.. (1997) (Correct)

This paper surveys the five service layers provided by the Stanford InfoBus: protocols for managing items and collections (DLIOP), metadata (SMA), search (STARTS), payment (UPAI), and rights and oblig... / ...the information and services on the Web. The Stanford InfoBus... / ...make use of the power of standard Web search engines such as Digital's AltaVista as soon...

#### 4 RAW: A Relational Algebra for the Web - Fiebig, Weiss, Moerkotte (1997) (Correct)

The main idea underlying the paper is to extend the relational algebra such that it becomes possible to process queries against the World-Wide Web. These extensions are minor in that we tried to keep ... / ... RAW A Relational Algebra for the Web Thorsten Fiebig Jurgen Weiss Guido ... / ...the functionality of a typical meta-search engine is capable of computing a unified...

#### 4 Knowledge Representation on the Web - Decker, Fensel, van Harmelen.. (2000) (Correct)

this paper, we make the following claims: unknown Knowledge Representation on the Web Stefan Decker, Dieter Fensel, Frank van Harmelen, Ian Horrocks, Sergey Melnik, Michel Klein and Jeen Broekstra ... / ... Knowledge Representation on the Web Stefan Decker Dieter Fensel ... / ...the use of ontologies can be found in search engines. By using ontologies the search...

#### 4 Finding near-replicas of documents on the web - Narayanan Shivakumar (1998) (Correct)

We consider how to efficiently compute the overlap between all pairs of web documents. This information can be used to improve web crawlers, web archivers and in the presentation of search results, ... / ...near-replicas of documents on the web Narayanan Shivakumar Hector... / ... . Improved ranking functions in search engines If documents A B and C are ...

#### 4 Adaptive Retrieval Agents: Internalizing Local Context and Scaling up .. - Filippo Menczer (1999) (Correct)

This paper focuses on two machine learning abstractions springing from ecological models: (i) evolutionary adaptation by local selection, and (ii) selective query expansion by internalization of env... / ...Local Context and Scaling up to the Web FILIPPO MENCZER AND RICHARD K.... / ...InfoSpiders could complement current search engine technology by starting up where...

#### 4 A Semi-Discrete Matrix Decomposition for Latent Semantic Indexing in... - Tamara Kolda (1997)

(Correct)

The vast amount of textual information available today is useless unless it can be effectively and efficiently searched. In information retrieval, we wish to match queries with relevant documents. Doc... / ... such as Alta Vista for the World Wide Web. Oftentimes users are searching for ... / ... string-matching tool in Unix and search engines such as Alta Vista for the World...

4 Tools and Approaches for Developing Data-Intensive Web Applications.. - Fraternali (1999)(Correct)

ions Implementation-level: pages, links, presentation styles Reuse Plug-in components; Reusable presentation styles Architecture Two-tiers, based on file system Static binding of content to pages ... / ...for Developing Data-Intensive Web Applications A Survey PIERO...

4 Design of Experiments in BDD Variable Ordering: Lessons Learned - Justin Harlow (1998)(Correct)

Applying the Design of Experiments methodology to the evaluation of BDD variable ordering algorithms has yielded a number of conclusive results. The methodology relies on the recently introduced equiv... / ...to popular search engines on the Web return up to and hits ... / ... processes keyword entries to popular search engines on the Web return up to and...

4 Equal Time for Data on the Internet with WebSemantics - George Mihaila (1998) (Correct)

Many collections of scientific data in particular disciplines are available today around the world. Much of this data conforms to some agreed upon standard for data exchange, i.e., a standard schema... / ...Time for Data on the Internet with WebSemantics George A. Mihaila ... / ... relevant documents through the use of search engines. Because of the inherent nature of...

4 Webcrawling Using Sketches - Lew, Lempinen, Huijsmans (1997) (Correct)

The current breed of WWW search engines are systems which use agents to find and download text based information from the distributed multimedia database known as the World Wide Web. These search engi... / ...final version of our paper entitled Webcrawling Using Sketches for... / ... Abstract The current breed of WWW search engines are systems which use agents to find ...

4 Research Issues in Web Data Mining - Madria, Bhowmick, Ng, Lim (1999) (Correct)

In this paper, we present an overview of research issues in web mining. We discuss mining with respect to web data referred here as web data mining. In particular, our focus is on web data mining ... / ... Research Issues in Web Data Mining SANJAY MADRIA SOURAV...

4 Searching the Web: General and Scientific Information Access - Lawrence, Giles (1999)(Correct)

The World Wide Web has revolutionized the way that people access information, and has opened up new possibilities in areas such as digital libraries, general and scientific information dissemination a... / ...Copyright c IEEE. Searching the Web General and Scientific Information...

ResearchIndex - [researchindex.org](http://researchindex.org) - [Terms of Service](#) - [Privacy Policy](#) - Copyright © 1997-2001 [NEC Research Institute](#)